RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/047,253A
Source: TFW16
Date Processed by STIC: 6/14/06

ENTERED

IFW16

RAW SEQUENCE LISTING DATE: 06/14/2006
PATENT APPLICATION: **US/10/047,253A** TIME: 10:18:13

Input Set : **A:\CIT1510-4.ST25.txt**
Output Set: **N:\CRF4\06142006\J047253A.raw**

```
      3 <110> APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
      4       COPE, Gregory
      5       VERMA, Rati
      6       ARAVIND, L.
      7       KOONIN, Eugene V.
      8       DESHAIES, Raymond
      9       AMBROGGIO, Xavier
     11 <120> TITLE OF INVENTION: REGULATION OF TARGET PROTEIN ACTIVITY THROUGH MODIFIER
PROTEINS
     13 <130> FILE REFERENCE: JHU1510-4
     15 <140> CURRENT APPLICATION NUMBER: US 10/047,253A
     16 <141> CURRENT FILING DATE: 2002-01-14
     18 <150> PRIOR APPLICATION NUMBER: US 60/261,314
     19 <151> PRIOR FILING DATE: 2001-01-12
     21 <150> PRIOR APPLICATION NUMBER: US 60/322,322
     22 <151> PRIOR FILING DATE: 2001-09-14
     24 <150> PRIOR APPLICATION NUMBER: US 60/322,030
     25 <151> PRIOR FILING DATE: 2001-09-14
     27 <160> NUMBER OF SEQ ID NOS: 24
     29 <170> SOFTWARE: PatentIn version 3.3
     31 <210> SEQ ID NO: 1
     32 <211> LENGTH: 14
     33 <212> TYPE: PRT
     34 <213> ORGANISM: Artificial sequence
     36 <220> FEATURE:
     37 <223> OTHER INFORMATION: JAM domain
     40 <220> FEATURE:
     41 <221> NAME/KEY: MISC_FEATURE
     42 <222> LOCATION: (1)..(14)
     43 <223> OTHER INFORMATION: Xaa is any amino acid
     45 <400> SEQUENCE: 1
W--> 47 His Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Asp
     48 1               5                   10
     51 <210> SEQ ID NO: 2
     52 <211> LENGTH: 17
     53 <212> TYPE: PRT
     54 <213> ORGANISM: Artificial sequence
     56 <220> FEATURE:
     57 <223> OTHER INFORMATION: JAM domain
     60 <220> FEATURE:
     61 <221> NAME/KEY: MISC_FEATURE
     62 <222> LOCATION: (3)..(3)
     63 <223> OTHER INFORMATION: Xaa is Tyr or Ile
     65 <220> FEATURE:
```

RAW SEQUENCE LISTING DATE: 06/14/2006
PATENT APPLICATION: **US/10/047,253A** TIME: 10:18:13

Input Set : **A:\CIT1510-4.ST25.txt**
Output Set: **N:\CRF4\06142006\J047253A.raw**

```
     66 <221> NAME/KEY: MISC_FEATURE
     67 <222> LOCATION: (5)..(5)
     68 <223> OTHER INFORMATION: Xaa is Ser or Thr
     70 <220> FEATURE:
     71 <221> NAME/KEY: MISC_FEATURE
     72 <222> LOCATION: (8)..(16)
     73 <223> OTHER INFORMATION: Xaa is any amino acid
     75 <400> SEQUENCE: 2
W--> 77 Gly Trp Xaa His Xaa His Pro Xaa Xaa Xaa Xaa Xaa Xaa Ser Xaa Xaa
     78 1               5                   10                  15
     81 Asp
     85 <210> SEQ ID NO: 3
     86 <211> LENGTH: 246
     87 <212> TYPE: PRT
     88 <213> ORGANISM: Homo sapiens
     90 <400> SEQUENCE: 3
     92 Thr Met Ile Ile Met Asp Ser Phe Ala Leu Pro Val Glu Gly Thr Glu
     93 1               5                   10                  15
     96 Thr Arg Val Asn Ala Gln Ala Ala Ala Tyr Glu Tyr Met Ala Ala Tyr
     97             20                  25                  30
    100 Ile Glu Asn Ala Lys Gln Val Gly Arg Leu Glu Asn Ala Ile Gly Trp
    101         35                  40                  45
    104 Tyr His Ser His Pro Gly Tyr Gly Cys Trp Leu Ser Gly Ile Asp Val
    105     50                  55                  60
    108 Ser Thr Gln Met Leu Asn Gln Gln Phe Gln Glu Pro Phe Val Ala Val
    109 65                  70                  75                  80
    112 Val Ile Asp Pro Thr Arg Thr Ile Ser Ala Gly Lys Val Asn Leu Gly
    113                 85                  90                  95
    116 Ala Phe Arg Thr Tyr Pro Lys Gly Tyr Lys Pro Pro Asp Glu Gly Pro
    117             100                 105                 110
    120 Ser Glu Tyr Gln Thr Ile Pro Leu Asn Lys Ile Glu Asp Phe Gly Val
    121         115                 120                 125
    124 His Cys Lys Gln Tyr Tyr Ala Leu Glu Val Ser Tyr Phe Lys Ser Ser
    125     130                 135                 140
    128 Leu Asp Arg Lys Leu Leu Glu Leu Leu Trp Asn Lys Tyr Trp Val Asn
    129 145                 150                 155                 160
    132 Thr Leu Ser Ser Ser Ser Leu Leu Thr Asn Ala Asp Tyr Thr Thr Gly
    133                 165                 170                 175
    136 Gln Val Phe Asp Leu Ser Glu Lys Leu Glu Gln Ser Glu Ala Gln Leu
    137             180                 185                 190
    140 Gly Arg Gly Ser Phe Met Leu Gly Leu Glu Thr His Asp Arg Lys Ser
    141         195                 200                 205
    144 Glu Asp Lys Leu Ala Lys Ala Thr Arg Asp Ser Cys Lys Thr Thr Ile
    145     210                 215                 220
    148 Glu Ala Ile His Gly Leu Met Ser Gln Val Ile Lys Asp Lys Leu Phe
    149 225                 230                 235                 240
    152 Asn Gln Ile Asn Ile Ser
    153                 245
    156 <210> SEQ ID NO: 4
```

RAW SEQUENCE LISTING DATE: 06/14/2006
PATENT APPLICATION: US/10/047,253A TIME: 10:18:13

Input Set : A:\CIT1510-4.ST25.txt
Output Set: N:\CRF4\06142006\J047253A.raw

```
157 <211> LENGTH: 245
158 <212> TYPE: PRT
159 <213> ORGANISM: Homo sapiens
161 <400> SEQUENCE: 4
163 Thr Val Arg Val Ile Asp Val Phe Ala Met Pro Gln Ser Gly Thr Gly
164 1               5                  10                  15
167 Val Ser Val Glu Ala Val Asp Pro Val Phe Gln Ala Lys Met Leu Asp
168             20                  25                  30
171 Met Leu Lys Gln Thr Gly Arg Pro Glu Met Val Val Gly Trp Tyr His
172         35                  40                  45
175 Ser His Pro Gly Phe Gly Cys Trp Leu Ser Gly Val Asp Ile Asn Thr
176     50                  55                  60
179 Gln Gln Ser Phe Glu Ala Leu Ser Glu Arg Ala Val Ala Val Val Val
180 65                  70                  75                  80
183 Asp Pro Ile Gln Ser Val Lys Gly Lys Val Val Ile Asp Ala Phe Arg
184                 85                  90                  95
187 Leu Ile Asn Ala Asn Met Met Val Leu Gly His Glu Pro Arg Gln Thr
188             100                 105                 110
191 Thr Ser Asn Leu Gly His Leu Asn Lys Pro Ser Ile Gln Ala Leu Ile
192         115                 120                 125
195 His Gly Leu Asn Arg His Tyr Tyr Ser Ile Thr Ile Asn Tyr Arg Lys
196     130                 135                 140
199 Asn Glu Leu Glu Gln Lys Met Leu Leu Asn Leu His Lys Lys Ser Trp
200 145                 150                 155                 160
203 Met Glu Gly Leu Thr Leu Gln Asp Tyr Ser Glu His Cys Lys His Asn
204                 165                 170                 175
207 Glu Ser Val Val Lys Glu Met Leu Glu Leu Ala Lys Asn Tyr Asn Lys
208             180                 185                 190
211 Ala Val Glu Glu Glu Asp Lys Met Thr Pro Glu Gln Leu Ala Ile Lys
212         195                 200                 205
215 Asn Val Gly Lys Gln Asp Pro Lys Arg His Leu Glu Glu His Val Asp
216     210                 215                 220
219 Val Leu Met Thr Ser Asn Ile Val Gln Cys Leu Ala Ala Met Leu Asp
220 225                 230                 235                 240
223 Thr Val Val Phe Lys
224                 245
227 <210> SEQ ID NO: 5
228 <211> LENGTH: 421
229 <212> TYPE: PRT
230 <213> ORGANISM: Homo sapiens
232 <400> SEQUENCE: 5
234 Met Pro Asp His Thr Asp Val Ser Leu Ser Pro Glu Glu Arg Val Arg
235 1               5                  10                  15
238 Ala Leu Ser Lys Leu Gly Cys Asn Ile Thr Ile Ser Glu Asp Ile Thr
239             20                  25                  30
242 Pro Arg Arg Tyr Phe Arg Ser Gly Val Glu Met Glu Arg Met Ala Ser
243         35                  40                  45
246 Val Tyr Leu Glu Glu Gly Asn Leu Glu Asn Ala Phe Val Leu Tyr Asn
247     50                  55                  60
```

RAW SEQUENCE LISTING DATE: 06/14/2006
PATENT APPLICATION: US/10/047,253A TIME: 10:18:13

Input Set : A:\CIT1510-4.ST25.txt
Output Set: N:\CRF4\06142006\J047253A.raw

```
250 Lys Phe Ile Thr Leu Phe Val Glu Lys Leu Pro Asn His Arg Asp Tyr
251 65                  70                  75                  80
254 Gln Gln Cys Ala Val Pro Glu Lys Gln Asp Ile Met Lys Lys Leu Lys
255                 85                  90                  95
258 Glu Ile Ala Phe Pro Arg Thr Asp Glu Leu Lys Asn Asp Leu Leu Lys
259             100                 105                 110
262 Lys Tyr Asn Val Glu Tyr Gln Glu Tyr Leu Gln Ser Lys Asn Lys Tyr
263         115                 120                 125
266 Lys Ala Glu Ile Leu Lys Lys Leu Glu His Gln Arg Leu Ile Glu Ala
267     130                 135                 140
270 Glu Arg Lys Arg Ile Ala Gln Met Arg Gln Gln Gln Leu Glu Ser Glu
271 145                 150                 155                 160
274 Gln Phe Leu Phe Phe Glu Asp Gln Leu Lys Lys Gln Glu Leu Ala Arg
275                 165                 170                 175
278 Gly Gln Met Arg Ser Gln Gln Thr Ser Gly Leu Ser Glu Gln Ile Asp
279             180                 185                 190
282 Gly Ser Ala Leu Ser Cys Phe Ser Thr His Gln Asn Asn Ser Leu Leu
283         195                 200                 205
286 Asn Val Phe Ala Asp Gln Pro Asn Lys Ser Asp Ala Thr Asn Tyr Ala
287     210                 215                 220
290 Ser His Ser Pro Pro Val Asn Arg Ala Leu Thr Pro Ala Ala Thr Leu
291 225                 230                 235                 240
294 Ser Ala Val Gln Asn Leu Val Val Glu Gly Leu Arg Cys Val Val Leu
295                 245                 250                 255
298 Pro Glu Asp Leu Cys His Lys Phe Leu Gln Leu Ala Glu Ser Asn Thr
299             260                 265                 270
302 Val Arg Gly Ile Glu Thr Cys Gly Ile Leu Cys Gly Lys Leu Thr His
303         275                 280                 285
306 Asn Glu Phe Thr Ile Thr His Val Ile Val Pro Lys Gln Ser Ala Gly
307     290                 295                 300
310 Pro Asp Tyr Cys Asp Met Glu Asn Val Glu Glu Leu Phe Asn Val Gln
311 305                 310                 315                 320
314 Asp Gln His Asp Leu Leu Thr Leu Gly Trp Ile His Thr His Pro Thr
315                 325                 330                 335
318 Gln Thr Ala Phe Leu Ser Ser Val Asp Leu His Thr His Cys Ser Tyr
319             340                 345                 350
322 Gln Leu Met Leu Pro Glu Ala Ile Ala Ile Val Cys Ser Pro Lys His
323         355                 360                 365
326 Lys Asp Thr Gly Ile Phe Arg Leu Thr Asn Ala Gly Met Leu Glu Val
327     370                 375                 380
330 Ser Ala Cys Lys Lys Lys Gly Phe His Pro His Thr Lys Glu Pro Arg
331 385                 390                 395                 400
334 Leu Phe Ser Ile Cys Lys His Val Leu Val Lys Asp Ile Lys Ile Ile
335                 405                 410                 415
338 Val Leu Asp Leu Arg
339             420
342 <210> SEQ ID NO: 6
343 <211> LENGTH: 461
344 <212> TYPE: PRT
```

RAW SEQUENCE LISTING DATE: 06/14/2006
PATENT APPLICATION: US/10/047,253A TIME: 10:18:13

Input Set : A:\CIT1510-4.ST25.txt
Output Set: N:\CRF4\06142006\J047253A.raw

345 <213> ORGANISM: Homo sapiens
347 <400> SEQUENCE: 6
349 Met Asp Gln Pro Phe Thr Val Asn Ser Leu Lys Lys Leu Ala Ala Met
350 1 5 10 15
353 Pro Asp His Thr Asp Val Ser Leu Ser Pro Glu Glu Arg Val Arg Ala
354 20 25 30
357 Leu Ser Lys Leu Gly Cys Asn Ile Thr Ile Ser Glu Asp Ile Thr Pro
358 35 40 45
361 Arg Arg Tyr Phe Arg Ser Gly Val Glu Met Glu Arg Met Ala Ser Val
362 50 55 60
365 Tyr Leu Glu Glu Gly Asn Leu Glu Asn Ala Phe Val Leu Tyr Asn Lys
366 65 70 75 80
369 Phe Ile Thr Leu Phe Val Glu Lys Leu Pro Asn His Arg Asp Tyr Gln
370 85 90 95
373 Gln Cys Ala Val Pro Glu Lys Gln Asp Ile Met Lys Lys Leu Lys Glu
374 100 105 110
377 Ile Ala Phe Pro Arg Thr Asp Glu Leu Lys Asn Asp Leu Leu Lys Lys
378 115 120 125
381 Tyr Asn Val Glu Tyr Gln Glu Tyr Leu Gln Ser Lys Asn Lys Tyr Lys
382 130 135 140
385 Ala Glu Ile Leu Lys Lys Leu Glu His Gln Arg Leu Ile Glu Ala Glu
386 145 150 155 160
389 Arg Lys Arg Ile Ala Gln Met Arg Gln Gln Gln Leu Glu Ser Glu Gln
390 165 170 175
393 Phe Leu Phe Phe Glu Asp Gln Leu Lys Lys Gln Glu Leu Ala Arg Gly
394 180 185 190
397 Gln Met Arg Ser Gln Gln Thr Ser Gly Leu Ser Glu Gln Ile Asp Gly
398 195 200 205
401 Ser Ala Leu Ser Cys Phe Ser Thr His Gln Asn Asn Ser Leu Leu Asn
402 210 215 220
405 Val Phe Ala Asp Gln Pro Asn Lys Ser Asp Ala Thr Asn Tyr Ala Ser
406 225 230 235 240
409 His Ser Pro Pro Val Asn Arg Ala Leu Thr Pro Ala Ala Thr Leu Ser
410 245 250 255
413 Ala Val Gln Asn Leu Val Val Glu Gly Leu Arg Cys Val Val Leu Pro
414 260 265 270
417 Glu Asp Leu Cys His Lys Phe Leu Gln Leu Ala Glu Ser Asn Thr Val
418 275 280 285
421 Arg Gly Ile Glu Thr Cys Gly Ile Leu Cys Gly Lys Leu Thr His Asn
422 290 295 300
425 Glu Phe Thr Ile Thr His Val Ile Val Pro Lys Gln Ser Ala Gly Pro
426 305 310 315 320
429 Asp Tyr Cys Asp Met Glu Asn Val Glu Glu Leu Phe Asn Val Gln Asp
430 325 330 335
433 Gln His Asp Leu Leu Thr Leu Gly Trp Ile His Thr His Pro Thr Gln
434 340 345 350
437 Thr Ala Phe Leu Ser Ser Val Asp Leu His Thr His Cys Ser Tyr Gln
438 355 360 365
441 Leu Met Leu Pro Glu Ala Ile Ala Ile Val Cys Ser Pro Lys His Lys

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 06/14/2006
PATENT APPLICATION: US/10/047,253A TIME: 10:18:14

Input Set : A:\CIT1510-4.ST25.txt
Output Set: N:\CRF4\06142006\J047253A.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; Xaa Pos. 2,4,5,6,7,8,9,10,11,12,13
Seq#:2; Xaa Pos. 3,5,8,9,10,11,12,13,15,16

VERIFICATION SUMMARY DATE: 06/14/2006
PATENT APPLICATION: **US/10/047,253A** TIME: 10:18:14

Input Set : **A:\CIT1510-4.ST25.txt**
Output Set: **N:\CRF4\06142006\J047253A.raw**

L:47 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:0
L:77 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:0